DOCUMENT RESUME

ED 279 678 TM 870 065

AUTHOR Guthrie, John T.
TITLE Roles of the National Assessment of Educational Progress in International Studies.
PUB DATE 23 Jul 86
NOTE 13p.; One of 46 papers commissioned by the Study Group on the National Assessment of Student Achievement and cited in Appendix B to their final report "The Nation's Report Card" (TM 870 049). For other papers in this group, see TM 870 050-094.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; Agency Cooperation; *Comparative Testing; Cross Cultural Studies; *Educational Assessment; Educational Testing; Elementary Secondary Education; Evaluation Utilization; Foreign Countries; *International Cooperation; Measurement Objectives; National Surveys; *Testing Programs; Test Use
IDENTIFIERS International Assn Evaluation Educ Achievement; *National Assessment of Educational Progress

ABSTRACT

This paper discusses the rationale for international comparisons in educational achievement, presents factors which affect the validity of such comparisons, and makes recommendations for the role of the National Assessment of Educational Progress (NAEP) with respect to international studies. International comparisons are of interest in policy formation, resource allocation, and school improvement programs. Four prerequisites to international assessment also weigh heavily in an individual, national testing program: (1) the scope of item content must be equitable for each of the countries involved; (2) comparisons across countries will be facilitated by the use of common scaling techniques; (3) sampling must be representative and aedquate; and (4) the appropriate language must be used in testing. NAEP can facilitate international comparisons by collaborating with existing efforts such as those of the International Association for the Evaluation of Educational Achievement (IEA). NAEP can provide technical assistance in data collection and statistical analysis. An official liaison between NAEP and IEA is recommended at the policy level. An annual cooperative working meeting is also recommended. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED279678

Roles of the National Assessment of Educational Progress in International Studies

John T. Guthrie

University of Maryland

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. T. Guthrie

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper commissioned by

THE STUDY GROUP ON THE NATIONAL ASSESSMENT OF STUDENT ACHIEVEMENT

1986

TM 870 065

BEST COPY AVAILABLE

ROLES OF THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS IN INTERNATIONAL STUDIES

A Report to

Department of Education

Task Force on Assessment Issues

Gov. Lamar Alexander, Chairman

Mr. H. Thomas James, Vice Chairman

from

John T. Guthrie

Center for Educational Research and Development

College of Education

University of Maryland

College Park, MD 20742

July 23, 1986

ROLES OF THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS IN INTERNATIONAL STUDIES

Purposes of This Paper

The purposes of this paper are to: 1) discuss the rationale for international comparisons in education; 2) present factors that affect the validity of such comparisons; and, 3) make recommendations for the roles of NAEP with respect to international studies.

History of International Comparisons

Educational research in which the outcomes of different nations are compared has been conducted by a number of organizations including: The Organization for Economic Cooperation and Development (OECD), The Council of Europe (COE), The United Nations Educational, Scientific and Cultural Organization (UNESCO) and The International Association for the Evaluation of Educational Achievement (IEA).

The most systematic and empirical of these organizations has been the IEA, which was begun by a group of educational researchers in the late 1950's. Its original purpose was to explore the effects of independent

variables on student performance in several school subjects across languages and across national boundaries. After a preliminary study of achievement in mathematics in 1966, IEA conducted a six-subject survey of performance in science, reading, literature, civic education, and two languages (French, and English) which was completed in 1974. Follow-up studies have been completed in mathematics, and science. By 1980 studies of the classroom environment and written composition has been initiated. In most cases, the international reports have presented general data about the comparative performance of education systems and have explained variations in level of performance in terms of variation in such areas as curriculum and instruction, school organization, community and family, and the personal characteristics of the students. Although the IEA does not collect periodic data from a constant set of countries, it represents the most well developed source of international comparisons involving USA and other countries.

Rationale for International Comparisons

There at least three reasons for conducting educational research at an international level. First, measures of educational achievement which provide

comparisons of school outcomes are of intrinsic interest to a wide range of consumers. Researchers, policy-makers, teachers, and the public are continually curious about how children in the United States compare to children in other countries. As the United States becomes more involved with other nations on economic, and cultural fronts, questions about comparative educational accomplishments increase. Indicators of comparative progress in education attract growing public and political attention as nations become more interdependent.

A second reason for conducting international comparisons in education is that policy-makers at the Federal, State and Local levels value them for determining the allocation of resources. The standing of the US relative to other countries in subjects such as reading, math and science, influence resources that are directed to these subjects. For example, recent comparisons of the US with technologically advanced nations such as Japan and West Germany have contributed to the sense of need for improvement in science education. Policy documents such as "A Nation at Risk" have used international comparisons, derived from IEA data, to make
the case that the United States is mediocre in its

performance. These data fueled the reform movement of 1983-1986.

The international data base on educational achievement is, regrettably, in a state of disrepair. A reporter from the Associated Press, with whom I was talking today, was aghast that the most recent data on reading was more than ten years old and appalled that it included only 15 countries. Since an international body such as the United Nations has not undertaken systematic, periodic comparisons of education, equal in quality to indicators of our economic health, it may be necessary for the USA through NAEP to exert leadership on this front.

A third reason for conducting comparisons of educational performance is the improvement of schools. Although NAEP has been devoted exclusively to data on school outcomes, data on inputs are also valuable. Frequently, the approach to school improvement following international comparisons is haphazard. A country that has performed poorly on an achievement measure will undertake an improvement program that is guided by fashion or expedience rather than by data. In addition, educators from countries that have performed well on an achievement

measure often make unwarranted claims about an aspect of the education system, such as the "teaching method" or the "materials" when those factors have not been the causal agent in producing achievement. Although the data from NAEP do not go beyond outcomes, the interpretations based on them nearly always extend to educational inputs. Such interpretations will be enhanced by information on the inputs of schooling.

When international comparisons are used for school improvement, data on educational inputs as well as educational outcomes are valuable. Types of inputs that are useful include: instructional strategies, school resources, student characteristics, educational materials, community context, and school policies. IEA has collected data on some of these variables in prior studies, but NAEP has not approached the issue of collecting data on indicators of educational inputs. Although it may be premature to undertake national surveys of inputs since measurements of outcomes have not been perfected, it is sensible to forecast that these factors will inform school improvement efforts and it is reasonable to design assessments of outcomes that are compatible with school improvement initiatives.

Requirements for International Comparisons

Comparing the educational achievement of students across countries is a demanding exercise. There are at least four prerequisites to international assessment some of which also weigh heavily in an individual, national testing program.

Scope. The content of the assessment items in an international study must be equitable for each of the countries involved. At the heart of the issue of scope is the opportunity for students to learn the skills and knowledge that are measured. Since it is known that achievement in sub-topics within math and reading are associated with time devoted to these topics in the curriculum, the adequacy with which the assessment represents the curriculum is an issue for many participating countries. The IEA has addressed this issue in the past by constructing an international core of tasks. All countries participating in a study are assessed on the core. In parallel, a description of the opportunities to learn all dimensions of the core is completed, permitting both absolute (unweighted) and relative (weighted) comparisons across countries.

Scales. Comparisons across countries will be facilitated by the use of common scaling techniques. Not only are equitable tests needed, but comparisons will benefit from scales that are constructed to describe the performance of participants in a common metric. The recent advances in scaling at the National Assessment of Educational Progress in the area of reading could be extended to an international data set. The technical requirements of this extension are not trivial however, and working meetings of content area specialists and psychometricians will be needed to address such issues as spiraling, weighted scoring, scale anchoring, and multidimensionality. Most of these problems can be resolved with existing statistical techniques, although their applications will require a diversity of expertise.

Sampling. Obtaining representative samples across countries is a challenging logistical operation. Most countries have nationalized data banks on schools, districts, pupils and personnel that permit stratified sampling. The size of the sample required for adequate representation varies with the intraclass correlation, which is the proportion of variance in achievement within a country that is attributable to differences between

schools. Countries in which differences between schools are large require larger samples of students to keep sampling error within tolerable limits.

Executing the plan is a problem. Obtaining a high rate of response to requests for participation in the study has been a more severe shortcoming in United States than other countries in prior IEA studies. NAEP could assist in addressing this issue. Consistently high response rates across countries are necessary to the creditability of international contrasts.

Linguistic. Assessments across languages are inevitably difficult. In tests of math as well as reading and science, cross translation is needed to assure that the coherence of the tasks and difficulty of vocabulary are similar in different languages. Furthermore, tests are usually administered in the language of the school, but this is not necessarily the language of the student's home and the penalty to bilingual populations must be addressed in the interpretation of findings. In multilingual or multicultural settings the administration of tests is complicated and special procedures are necessary to assure that students have understood the tasks to be performed,

and have attached appropriate importance to the assessment activity.

Facilitation of International Comparisons

Comparisons of educational achievement that will embrace the United States and other countries can be facilitated by the National Assessment of Educational Progress at least three ways. First, NAEP can collaborate with international organizations that are collecting data on educational systems. The most prominent of these organizations is the International Association for the Evaluation of Educational Achievement (IEA) which has been founded to foster such studies.

The US participation in IEA studies in such areas as reading, math and science should be directed by committees of national experts in the respective content domains. The NAEP could be contracted to provide services such as data collection, statistical analysis and consultation in accordance with the design of the national and international committees that direct the studies. Through this structure NAEP can lend technical expertise to the policy and research priorities shared by the USA and other countries in the IEA organization.

Second, an official liaison between NAEP and IEA

should be established at the policy level. A member of the NAEP policy board could work with administrative officers of IEA to construct concrete lines of mutual support. The NAEP policy board could invite an officer of the IEA to serve as a voting or an ex-officio member. A meeting of minds at the policy level is necessary.

Third, to foster communication, it is recommended that NAEP host a select, invited, working meeting on an annual basis for exchange between NAEP and IEA research groups. The areas of reading, literacy, math and science could be represented at a minimum. The design of assessments, sampling schemes, item banks, spiraling or matrix sampling procedures, data analysis, and archival strategies must be shared more broadly than they have in past.

An exploratory meeting was held on May 6, 1986 with NAEP reading and literacy staff, OERI staff, and the IEA Literacy Steering Committee. After an exchange of general assessment designs and plans, the group concurred that joint data collection, analysis, reporting and publishing would be mutually beneficial. Continuing the flow of communication with the international research community is a first step to enhancement of the role of NAEP in educational research on international level.

JTG/md

7/28/86